

# Identification of QTLs for sex expression in dioecious and monoecious hemp (*Cannabis sativa* L.)

A.-M. Faux · X. Draye · M.-C. Flamand ·  
A. Occre · P. Bertin

Received: 28 August 2015 / Accepted: 10 January 2016 / Published online: 25 January 2016  
© Springer Science+Business Media Dordrecht 2016

**Abstract** Hemp (*Cannabis sativa* L.) is a diploid species including both dioecious and monoecious cultivars with hetero- and homomorphic sex chromosomes, respectively. It displays a high plasticity of sex expression, i.e., the ratio of female and male flowers. In this study, we investigated the role of sex chromosomes in the genetic determinism of sex expression in dioecious and monoecious hemp. The experimental materials were three F<sub>1</sub> segregating populations, two dioecious (C1 and C2: ‘Carmagnola’ ♀ × ‘Carmagnola’ ♂), and one monoecious (UF: ‘Uso 31’ × ‘Fedora 17’). A ‘sex’ phenotypic marker was mapped in C1 and C2. In total, 23, 42, and 26 AFLP markers (71 markers in total) were mapped to three, nine, and three co-segregation groups putatively located on sex chromosomes in C1, C2, and UF, respectively. Recombination rates with sex ranged from 0 to 0.5. Five sex-linked markers were detected in UF, revealing

homologies between the X chromosomes of monoecious hemp and the X and Y chromosomes of dioecious hemp. Five QTLs associated with quantitative variations in sex expression were identified in each map. Four markers associated with variations in sex expression in UF segregated with sex or accounted for a putative QTL in C1 or C2. Two QTLs and three of these markers were mapped in UF in a region homologous to the sex-locus region of the dioecious maps. Given these results, conducting further research on the genetic determinism of sex expression in hemp using a quantitative approach appears relevant.

**Keywords** *Cannabis* · Dioecious · Monoecious · QTL · Sex chromosomes · Sex expression

## Introduction

Hemp, *Cannabis sativa* L., is a non-food crop currently grown for a wide range of end products derived from its fibre, wooden core, or seeds, such as textiles, paper pulp, lime-hemp concretes, or hempseed oil (Struik et al. 2000). After a decline in its cultivation over the last 200 years, this species is now perceived as a potentially profitable crop suitable for cultivation in sustainable farming systems (van der Werf et al. 1996).

Hemp is a diploid species ( $2n = 20$ ) that has heteromorphic sex chromosomes and is characterized

---

A.-M. Faux (✉) · X. Draye · A. Occre · P. Bertin  
Earth and Life Institute, Université catholique de Louvain,  
Croix du Sud 2 – 11 (L7.05.11), 1348 Louvain-la-Neuve,  
Belgium  
e-mail: am.faux@gmail.com

P. Bertin  
e-mail: pierre.bertin@uclouvain.be

M.-C. Flamand  
Institut des Sciences de la Vie, Université catholique de  
Louvain, Croix du Sud 4 – 5 (L7.07.14),  
1348 Louvain-la-Neuve, Belgium

by sexual dimorphism. Male plants of dioecious hemp have XY sex chromosomes, and female plants have XX sex chromosomes (Yamada 1943, cited by Sakamoto et al. 1995). The sex of the plant (male/female) is controlled by an X-to-autosomes equilibrium rather than by a Y-active system (Westergaard 1958; Ainsworth 2000). The presence of a pseudoautosomal region allowing recombination between the X and Y chromosomes has been reported in dioecious hemp (Peil et al. 2003), as well as the existence of polymorphisms between the X chromosomes (Peil et al. 2003; Rode et al. 2005). Morphologically, male plants of dioecious hemp are characterised by hanging panicles with few or no leaves, and staminate flowers. The female plants have a distinct appearance, bearing racemes with leafy bracts and pistillate flowers. Despite the presence of sex chromosomes, some plants of dioecious hemp produce flowers of the sex opposite to the main one. This anomalous flowering trait is affected by external factors, such as the photoperiod or hormonal treatments (Freeman et al. 1980).

Monoecious hemp has homomorphic sex chromosomes (XX) (Menzel 1964; Faux et al. 2014; Razumova et al. 2015). However, sex expression, defined as the ratio of female to male flowers, is highly variable, and the genetic determinism of this variation is unknown. The plants of monoecious hemp have inflorescences similar to those of female plants. However, they produce male and female flowers with ratios that differ among nodes along the stem, plants, cultivars, over time (Faux et al. 2013, 2014; Faux and Bertin 2014), and in response to external factors (Arnoux et al. 1966a, b; Arnoux and Mathieu 1969; Truta et al. 2007). The quantitative variation in sex expression along the stem of monoecious hemp plants has been successfully described as a logistic function of the node position (Faux and Bertin 2014). The observation of genotypic variability in sex expression in monoecious hemp led us to assume that this trait can be inherited, and therefore, it could be investigated by searching for quantitative trait loci (QTL) (Faux et al. 2014).

Agriculturally, hemp cultivation is significantly affected by the reproductive features of the plant. Although this species is naturally dioecious, both dioecious and monoecious cultivars exist. Dioecious cultivars have relatively high stem yields. However, male plants flower and senesce earlier than female

plants, so it is difficult to determine the optimal time for harvesting the fiber of such cultivars as well as to harvest the seeds mechanically. This restricts seed multiplication for such cultivars (Bocsa and Karus 1998). In contrast, monoecious cultivars mature synchronously, allowing for the mechanical harvest of both stems and seeds (Mandolino and Carboni 2004). Consequently, seeds of these cultivars are easily obtained. In addition, these cultivars can improve the profitability of the farming operation by allowing the production of both stems and seeds (Bocsa and Karus 1998).

In monoecious hemp cultivars, the masculinised plants set few seeds and are more prone to fungal infections (Fournier and Beherec 2006). Therefore, strongly masculinised plants must be eliminated to produce good-quality seeds from monoecious cultivars (Beherec 2000). A previous study showed that sex expression in monoecious hemp varied among cultivars, while higher seed yields were obtained from early and mid-early feminised cultivars (Faux et al. 2013). Given its practical implications in both cultivation and breeding, sex expression is considered to be one of the most important factors in the genetic improvement in hemp (Mandolino and Carboni 2004; Ranalli 2004).

So far, two non-saturated molecular maps have been published for hemp (Mandolino and Ranalli 2002; Peil et al. 2003, cited by Mandolino and Carboni 2004). Both maps were constructed using  $F_1$  segregating progenies. The first map (Mandolino and Ranalli 2002) was based on 40 individuals derived from a cross between a monoecious plant and a female plant; thus, the Y chromosome was not included. In that study, 66 RAPD markers segregating 1:1 were distributed among 11 co-segregation groups in the female parent map, and 43 markers were distributed among nine co-segregation groups in the monoecious parent map. The second map was obtained from a cross between a male plant and a female plant of a dioecious accession. A total of 122 AFLP markers was distributed among 10 linkage groups (Peil et al. 2003, cited by Mandolino and Carboni 2004).

Sex-specific AFLP markers were successfully detected in hemp by Flachowsky et al. (2001). The AFLP technique was also used to detect sex-specific markers in *Dioscorea tokoro* (Terauchi and Kahl 1999) and *Asparagus officinalis* (Reamon-Büttner and Jung 2000). The AFLP technique is known for its

reliability and the potentially high number of amplified fragments, which makes it a useful tool for the construction of genetic maps (Vos et al. 1995; Mueller and Wolfenbarger 1999; Meudt and Clarke 2007). Besides, according to Rouppe van der Voort et al. (1997) and Qi et al. (1998), the homology of AFLP products among populations belonging to the same species is nearly always valid, allowing AFLP markers to be used to integrate different linkage maps of a given species.

In this study, we investigated the roles of sex chromosomes in the genetic determinism of sex expression in dioecious and monoecious hemp. The experimental materials were segregating populations derived from dioecious and monoecious hemp cultivars. The use of a dioecious population allowed us to map a ‘sex’ phenotypic marker, and thus, to identify sex-linked markers. Genetic mapping was carried out using AFLP markers, and the genetic determinism of sex expression was investigated using a QTL analysis. We searched for QTLs using mixed models, because of their high flexibility to incorporate co-variables (such as time) into analyses, and to adequately model the residual genetic variation (Malosetti et al. 2004, 2006, 2007; Boer et al. 2007; Pastina et al. 2012). There were two specific objectives: (i) to identify sex-linked markers in dioecious hemp and, by homology, markers and co-segregation groups putatively located on the sex chromosomes in monoecious hemp; and (ii) to identify QTLs associated with sex expression in hemp.

## Materials and methods

### Genetic material

Three hemp cultivars were used: one dioecious, ‘Carmagnola’, and two monoecious, ‘Usó 31’ and ‘Fedora 17’. Seeds of the dioecious cultivar were obtained from Assocanapa (Carmagnola, Italy), and seeds of both monoecious cultivars were obtained from the Fédération Nationale des Producteurs de Chanvre (Le Mans, France). Three populations were created, all consisting of the full-sib  $F_1$  progeny of a cross between two outbred parents. The first two populations (C1 and C2) were derived from a cross between male and female ‘Carmagnola’ plants. The third population (UF) was obtained from a cross

between ‘Usó 31’ as the male parent and ‘Fedora 17’ as the female parent after emasculation. ‘Usó 31’ has been described as being more masculinised than ‘Fedora 17’, while both cultivars exhibit similar earliness, allowing their flowering periods to synchronise (Faux et al. 2013, 2014). The C1, C2, and UF populations included 77 (43 females and 34 males), 76 (48 females and 28 males), and 167 individuals, respectively. The relatively poor seed set in the controlled crosses in ‘Carmagnola’ made the use of two populations necessary.

### Growth conditions

The plants were cultivated in a greenhouse. One single plant per genotype was grown. Growth conditions were as described in Faux and Bertin (2014), except for the photoperiod, which was 16 h during the entire trial period for UF and 16 h for 69 days and 8 h thereafter to promote flowering of the C1 and C2 populations.

### Molecular data

Molecular data were extracted for each of the 320  $F_1$  individuals and six parental individuals. DNA extraction was conducted as described in Faux et al. (2014). AFLP amplifications were performed according to Vos et al. (1995) with slight modifications according to Flachowsky et al. (2001) and Peil (pers. comm. 2009). DNA was cleaved using the restriction enzymes HindIII and Tru91 (same restriction site as MseI) (Promega Corp, Madison, WI, USA). After ligation with adaptors specific to each restriction enzyme, pre-selective amplifications were performed with both HindIII- and Tru91-primers with one selective nucleotide (A). The pre-amplification mixture was diluted 20 times with sterile  $H_2O$ . The selective amplifications were performed using a HindIII-primer as a rare cutter and both HindIII- and Tru91-primers with three selective nucleotides. Eight distinct primer combinations (Flachowsky et al. 2001) were tested (Table 1). All primer and adapter sequences were designed by Eurogentec SA (Seraing, Belgium). PCRs were run on a PTC 100 thermal cycler (MJ Research, Waltham, MA, USA).

The amplification products were fractionated with an automated ABI Prism 3100 Genetic analyser (Applied Biosystems, Warrington, UK) according to

**Table 1** Selective nucleotides of each primer combination used to generate AFLP markers

No. <sup>a</sup>	Primer combination	
	<i>Hind</i> III	<i>Mse</i> I
1	AGA	AAC
2	AGA	AGG
3	AGA	AAG
4	ACC	AAG
5	ACC	AGA
6	ACC	ATC
7	ACT	ACA
8	ACT	ACC

<sup>a</sup> Number of primer combination used to label the markers

the manufacturer's instructions. Their sizes were scaled with the molecular standard GeneScan-500 Rox. The size and intensity of the amplified fragments were visualised using Peak Scanner™ v1.0 free software (Applied Biosystems 2006), and the fragments were between 40- and 400-bp long.

#### Segregation of AFLP markers

Each amplification was scored as a dominant marker—i.e., according to its presence (allele *a*) or absence (allele *o*)—in each individual. The markers were labelled with the number of the primer combination (Table 1) followed by the molecular weight of the DNA fragment (in bp). A cross type was attributed to each marker in each population according to its presence in the parents and the segregation pattern in the offspring, which was tested against both possible ratios (1:1 and 3:1) using Chi square tests ( $\chi^2$ ). Three cross types were distinguished: markers that were heterozygous in the female parent and absent from the male parent; markers that were heterozygous in the male parent and absent from the female parent (all of these segregating 1:1 in the offspring); and markers that were heterozygous in both parents, segregating 3:1.

#### Identification of sex-linked markers

In the dioecious populations, the sex of each plant was coded as 0 (female) or 1 (male) and used as a phenotypic marker. Sex-linked AFLP markers were detected using the five classes of markers segregating with the 'sex' phenotypic marker as defined by Peil

**Table 2** Five classes of sex-linked markers as defined by Peil et al. (2003)

Class <sup>a</sup>	Parents		Expected segregation ratios in F <sub>1</sub>	
	X    X <sup>b</sup>	X    Y	♀ (a:o)	♂ (a:o)
A	a    o	a    o	1:0	1:1
B	a    o	o    o	1:1	1:1
C	o    o	a    o	1:0	0:1
D	a    o	o    a	1:1	1:0
E	o    o	o    a	0:1	1:0

<sup>a</sup> Class B markers are detected by linkage to class A markers in male progenies and to class D markers in female progenies, in addition to falling into the expected cross-type and segregation ratios

<sup>b</sup> Location of alleles *a* and *o* is interchangeable between the two X chromosomes in the female parent

et al. (2003) (Table 2). Classes A, B, and C contain markers that are polymorphic on the X chromosome with no fragment on the Y chromosome. Class A markers are heterozygous in both parents, class B markers are heterozygous in the female parent only, and class C markers are heterozygous in the male parent only. Class D markers are heterozygous in both parents with the 'presence' allele located on the Y chromosome in the male parent, indicating a pseudoautosomal region. Class E includes male-associated markers: they are heterozygous in the male parent only, with the 'presence' allele located on the Y chromosome. Recombinants between class B and sex cannot be detected as a result of their respective allelic configurations. Therefore, class B markers are identified by linkage to class A markers in the male progenies and to class D markers in the female progenies. In the present study, the markers that responded to the cross type and segregation criteria of class B (Table 2) but were not linked to a class A or D were referred to as 'putative class B' markers. These markers could be located on both sex chromosomes and autosomes.

The markers were assigned to a class of sex-linked markers by testing their segregation ratios in male and female progenies against the expected segregation ratios using a Chi square test. Deviations from the expected segregation ratios at alpha = 0.01 were discarded.

## Map construction

The linkage analysis and map construction were performed independently for each segregating population using the OneMap package (Margarido et al. 2007, 2012). The linkage analysis was performed using the same statistical stringency to construct each of the three maps, to allow for meaningful comparisons among them. The LOD score and recombination fraction thresholds used were 3.5 and 0.5, respectively. The same threshold values of LOD score and recombination fraction were used to order markers within co-segregation groups. Map distances were based on the Kosambi function. The dioecious maps (C1 and C2) were integrated by using the markers that were mapped in both populations. Similarly, the monoecious map was integrated using the markers shared with C1 or C2.

## Phenotyping

Sex expression was characterized based on the reproductive morphology of the plants (dioecious or monoecious). In the dioecious populations, the plants were phenotyped at three times, i.e., 77, 84, and 98 days after sowing. The sex of each plant was recorded as a binary variable, i.e., male or female. However, flowers of the sex opposite to the main one were observed on some plant nodes. For each male plant at each observation time, %*F*-♂ indicated the percentage of nodes bearing female flowers. Similarly, for each female plant at each observation time, %*M*-♀ indicated the percentage of nodes bearing male flowers.

In the monoecious population, the continuous variation of sex expression was characterized by the monoecy degree (MD), a five-point scale ranging from 1 (mostly male flowers) to 5 (mostly female flowers) (Sengbusch 1952; Faux and Bertin 2014). We scored MD for each flowering node at six times, i.e., 43, 50, 57, 64, 71, and 78 days after sowing. Eight phenotypic variables were used to summarise the sex expression of each monoecious plant at each observation time (Table 3), as described in Faux and Bertin (2014).

## QTL analysis

The QTL analysis was performed in each segregating population independently. Putative QTLs were identified by interval mapping (Lander and Botstein 1989)

before being tested in a multiple regression similar to that applied by Pastina et al. (2012) to an *F*<sub>1</sub> segregating population of sugarcane.

### (i) Computation of genetic predictors

Three QTL effects were estimated: two additive effects,  $\alpha_P$  in parent *P* and  $\alpha_Q$  in parent *Q*, and one dominance effect,  $\delta_{PQ}$ . Genetic predictors were constructed for each of the three QTL effects (Lynch and Walsh 1998; Pastina et al. 2012) and computed for a grid of evaluation points along the genome. These predictors were introduced as explanatory variables in the QTL models (see below).

### (ii) Identification of putative QTLs by interval mapping (IM)

For notation purposes, the random variables are underlined in the equation below. The presence of a putative QTL at evaluation genomic point *w* was assessed using the following model, hereafter referred to as the IM model:

$$\underline{y}_i = \mu + x_{P_w} \alpha_{P_w} + x_{Q_w} \alpha_{Q_w} + x_{PQ_w} \delta_{PQ_w} + \underline{e}_i \quad [1]$$

where  $\underline{y}_i$  is the phenotypic value of individual *i*,  $\mu$  is the overall mean phenotypic value,  $\alpha_{P_w}$ ,  $\alpha_{Q_w}$  and  $\delta_{PQ_w}$  are the effects of each of the three genetic predictors,  $x_{P_w}$ ,  $x_{Q_w}$  and  $x_{PQ_w}$ , respectively, and  $\underline{e}_i$  is the residual error associated with individual *i*, normally distributed with mean 0 and variance  $\sigma^2$ .

For the binary variable sex, the IM model was tested using the GLIMMIX procedure with binomial distribution for the response variable and logit link function in SAS (SAS Institute Inc. 2012). All other phenotypic variables (Table 3) were analysed by allowing non-null correlations between the observations made at distinct times on the same plant. For this purpose, the MIXED procedure in SAS was used with the REPEATED statement including type = AR(1) and subject = individual.

The search for putative QTLs was performed for each phenotypic variable similar to the analysis described by Pastina et al. (2012). A putative QTL was considered to be found at position *w* if  $\alpha_P$ ,  $\alpha_Q$ , or  $\delta_{PQ}$  was significantly different from 0 at  $P < 0.01$ . The putative QTLs corresponding to a local minimum *p*-value were retained for the multi-QTL analysis. For the variable sex only, the percentage of variation ( $r^2$ )

**Table 3** Phenotypic variables used to characterize sex expression in dioecious and monoecious hemp populations

Population	Type <sup>a</sup>	Variable <sup>b</sup>	Description <sup>c</sup>
Dioecious	–	sex	Sex of plant, male or female
♂ plants		%F-♂	% of nodes bearing female flowers on male plants
♀ plants		%M-♀	% of nodes bearing male flowers on female plants
Monoecious	Synthesis	mMD	Mean MD
		%MD1	% of nodes with MD = 1
		%MD5	% of nodes with MD = 5
		%MDinter	% of nodes with 1 < MD < 5
	Structure	log_k	Logistic-function parameter determining curvature of the function
		log_NDm	Logistic-function parameter accounting for node at which variation of MD along the stem is maximum
		log_Mmin	Logistic-function parameter accounting for minimum MD along stem
	log_Mmax	Logistic-function parameter accounting for maximum MD along stem	

For details of the construction of variables used to characterize sex expression of monoecious plants, see Faux and Bertin (2014)

<sup>a</sup> In the monoecious population, structure variables consisted of parameters of a logistic curve describing variation of the monoecy degree as a function of node position along the stem

<sup>b</sup> Except for sex, all variables were longitudinal, i.e., their value was determined at several times

<sup>c</sup> MD = monoecy degree. For each plant, monoecy degree was scored for each flowering node at six observation times

explained by each additive effect was obtained from the fit of the IM model. This was not possible for the other variables because different methods were used to fit the model 1—i.e., restricted maximum likelihood rather than least-squares (SAS Institute Inc. 2012).

### (iii) Identification of QTLs by multi-QTL analysis

For each phenotypic variable, the genetic effects of each putative QTL as detected by the IM model were tested together in an additive multiple regression using the following model:

$$y_i = \mu + \sum_{w \in (W_P, W_{PQ})} x_{P_w} \alpha_{P_w} + \sum_{w \in (W_Q, W_{PQ})} x_{Q_w} \alpha_{Q_w} + \sum_{w \in W_{PQ}} x_{PQ_w} \delta_{PQ_w} + \epsilon_i \quad [2]$$

where  $W_P$ ,  $W_Q$ , and  $W_{PQ}$  are the sets of positions along the genome accounting for a putative QTL with an  $\alpha_P$ ,  $\alpha_Q$ , and  $\delta_{PQ}$  effect, respectively. The genetic predictors of both additive effects at the positions accounting for a putative QTL with a dominance effect ( $W_{PQ}$ ) were integrated into model 2 for more powerful QTL detection (Pastina et al. 2012). The multi-QTL analysis was performed using the GLIMMIX and MIXED procedures in SAS similar to the QTL analysis using the IM model (model 1).

A multi-QTL model was constructed for each phenotypic variable through a procedure of backward selection of genetic effects starting from model 2. At each step, model 2 was tested, and the effect associated with the highest non-significant ( $P > 0.05$ )  $p$ -value was removed until no effect was associated with a non-significant  $p$ -value. The resulting model was referred to as the multi-QTL model. The genomic positions selected in the multi-QTL model were considered to be QTLs, and the effect and standard error of each QTL were estimated from the multi-QTL model (Malosetti et al. 2006).

### Cross-populations analysis

A cross-population (CP) analysis was conducted to assess the consistency of the effect of putative QTLs across populations. This analysis was restricted to the putative QTLs mapped at marker positions, since these were the only positions that could be assumed to be homologous between different linkage maps. For this purpose, all of the markers accounting for a putative QTL as detected by the IM model (model 1) in at least one of the three populations were retained.

The CP analysis was first performed across both dioecious populations including a ‘population’ cofactor, and then across the dioecious and monoecious

populations. The consistency of the segregation pattern of the retained markers with sex across both dioecious populations was tested as follows:

$$\underline{y}_{ik} = \mu + \gamma_k + (\beta + \gamma\beta_k) \times M + \underline{e}_{ik} \quad [3]$$

where  $\underline{y}_{ik}$  is the sex of individual  $i$  from dioecious population  $k$  (male/female),  $\mu$  is the mean value over the two populations,  $\gamma_k$  is the effect of population  $k$ ,  $\beta$  is the mean effect of the marker, and  $(\beta\gamma)_k$  its effect in population  $k$ ,  $M$  is  $-1$  or  $1$  according to the presence/absence of the marker in the  $i$ th individual of population  $k$ , and  $\underline{e}_{ik}$  is the residual error associated with individual  $i$  in population  $k$ . Then, the effect of the retained markers was tested on the phenotypic variables recorded in the monoecious population using model 3 without the ‘population’ cofactor. The CP analysis was performed using the GLIMMIX and MIXED procedures in SAS similar to the QTL analysis using the IM model (model 1).

## Results

### AFLP markers

The eight primer combinations (Table 1) allowed the detection of 410, 415 and 357 markers in the C1, C2 and UF populations, respectively, generating a total of 480 distinct AFLP markers (Table 4). Out of these, 287 (60 %) were scored in all three populations. The percentage of markers shared by C1 and C2 (81 %) was higher than that shared by C1 and UF (67 %) or C2 and UF (68 %).

Among the 480 detected markers, 274 (67 % of 410 markers), 182 (44 % of 415 markers) and 184 (52 % of 357 markers) segregated in C1, C2 and UF, respectively. They accounted for a total of 385 distinct segregating markers, out of which 126 segregated in both C1 and C2, 111 in both C1 and UF, 72 in both C2 and UF, and 54 in all three populations.

Among the 385 segregating markers, 193 (70 % of 274 markers), 147 (81 % of 182 markers) and 115 (63 % of 184 markers) segregated independently (1:1 or 3:1) in C1, C2 and UF, respectively. These accounted for a total of 319 distinct markers available for mapping, out of which 75 were found in both C1 and C2, 46 in both C1 and UF, 34 in both C2 and UF, and 19 in all three populations. The independently segregating markers consisted of 71, 53 and 64

markers segregating 1:1 (heterozygous in only one of both parents), and 122 (63 % of 193 markers), 94 (64 % of 147 markers), and 51 (44 % of 115 markers) markers segregating 3:1 (heterozygous in both parents) in C1, C2, and UF, respectively.

### Linkage maps

The linkage analysis resulted in the mapping of 93, 92 and 86 AFLP markers assigned to 11, 16 and 10 co-segregation groups in C1, C2 and UF, respectively (not shown). These markers accounted for a total of 225 distinct mapped markers, among which 22 were shared by C1 and C2, the two ‘Carmagnola’ maps, 12 by C1 and UF, 17 by C2 and UF, and only five by all three maps. Out of these markers, 56, 67, and 41 % were heterozygous in both parents in each map, respectively (Table 4).

Among the mapped markers, 23, 42, and 26 were linked, directly or through homologous markers, to sex in C1, C2, and UF, respectively. They were distributed along 3, 6, and 3 co-segregation groups putatively located on sex chromosomes in each map, respectively (Fig. 1). They accounted for a total of 71 distinct markers, among which 9 were shared by C1 and C2, 3 by C1 and UF, 9 by C2 and UF, and a single one by all three maps (6\_255).

The C1, C2, and UF maps of the sex chromosomes covered total distances of 313.5, 399.8, and 489.2 cM, respectively (Fig. 1). The size of the CGs ranged between 6.1 and 252.9 cM in C1, 5.0 and 166.3 cM in C2, and 17.1 and 308.1 cM in UF. Markers were separated by an average distance of 13.7, 9.5, and 18.8 cM in the C1, C2, and UF maps, respectively, and a maximum distance of 43.5 (C1-2), 41.3 (C2-5), and 65.0 cM (UF-1) in each map, respectively. The largest gap (65.0 cM in UF-1) resulted from ignoring the recombination fraction between markers heterozygous in the male parent and those heterozygous in the female parent, as demonstrated by Maliepaard et al. (1997).

Co-segregation groups including the ‘sex’ phenotypic marker in the dioecious maps

The ‘sex’ phenotypic marker was mapped in C1-1 and C2-1 (Table 5; Fig. 1). These CGs, referred to as ‘sex CGs’, included eight and 19 AFLP markers, respectively—21 different markers in total. Among them, six

**Table 4** The number of markers scored, selected for mapping, and mapped in each population

	C1	C2	UF	All populations <sup>a</sup>
All scored markers	410	415	357	480
Segregating markers <sup>b</sup>	274 (67)	182 (44)	184 (52)	385
All independently segregating markers (1:1 or 3:1) <sup>c</sup>	193 (70)	147 (81)	115 (63)	319
Heterozygous in only one parent (segregation 1:1; “ <i>ao</i> × <i>oo</i> ” or “ <i>oo</i> × <i>ao</i> ”)	71	53	64	
Heterozygous in both parents (segregation 3:1; “ <i>ao</i> × <i>ao</i> ”) <sup>d</sup>	122 (63)	94 (64)	51 (44)	
All mapped markers	93	92	86	225
Heterozygous in only one parent	41	30	51	
Heterozygous in both parents <sup>d</sup>	52 (56)	62 (67)	35 (41)	
Markers mapped in putative sex co-segregation groups	23	42	26	71
Heterozygous in only one parent	18	16	17	
Heterozygous in both parents <sup>d</sup>	5 (22)	26 (62)	9 (35)	

<sup>a</sup> Total number of distinct markers regardless of the population

<sup>b</sup> The segregating markers had a segregating pattern significantly different from 1:0 (i.e., marker always present) as tested by a Chi<sup>2</sup> test with Bonferroni correction at alpha = 0.05. In parentheses, % of segregating markers in relation to the total number of scored markers

<sup>c</sup> In parentheses, % of independently segregating markers in relation to the total number of segregating markers

<sup>d</sup> In parentheses, % of markers heterozygous in both parents in relation to the total number of independently segregating markers, total number of mapped markers, and total number of markers mapped in putative sex CGs, respectively

and 12 markers—14 different markers—were assigned to classes A, D, or E (Table 2) in C1 and C2, respectively.

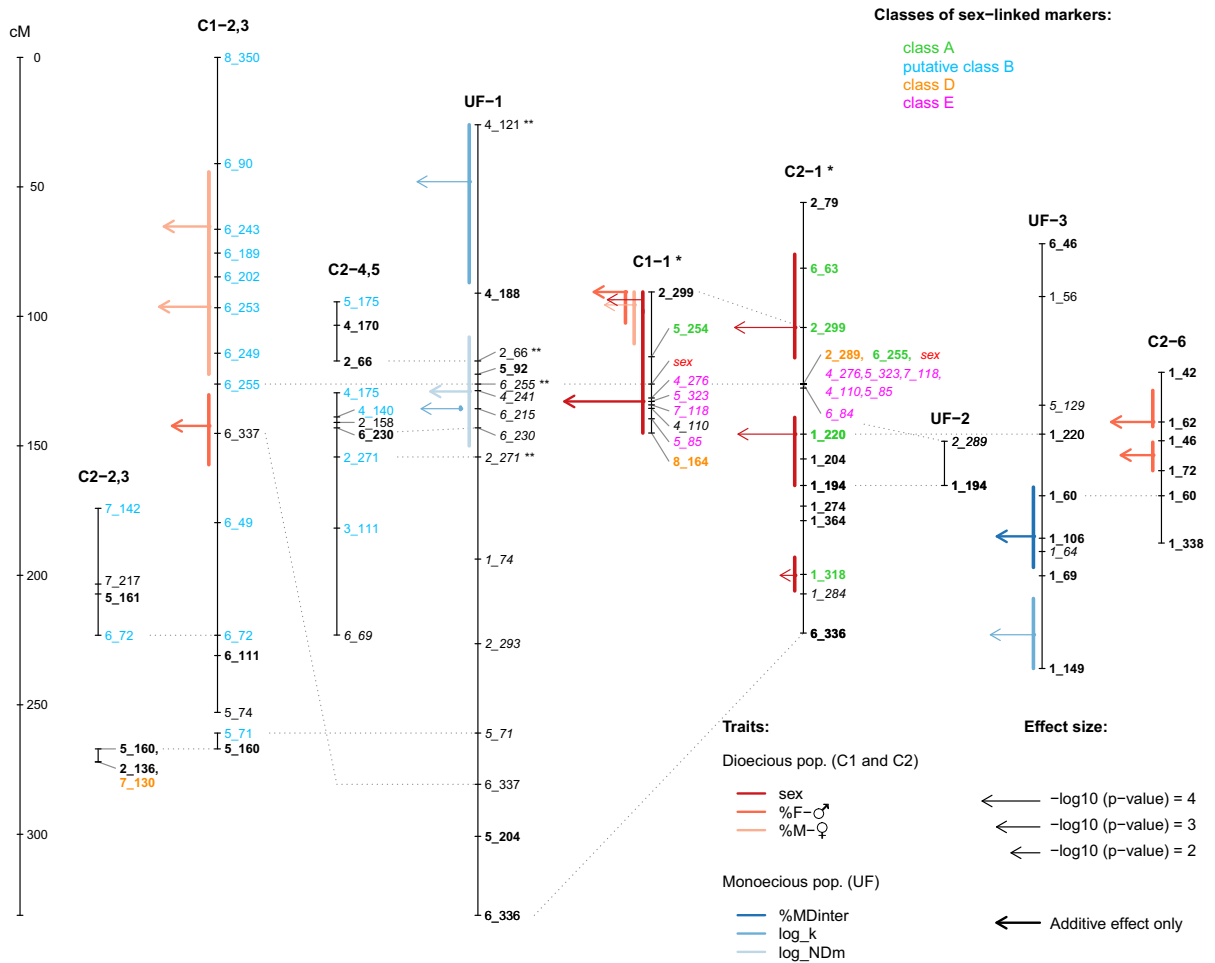
Four successive regions were determined in the sex CGs based on the recombination rate of the markers with sex (Table 5): (1) a region including class A markers in addition to markers that were not assigned to a given class of sex-linked markers; (2) a central region surrounding the sex locus and including all of the class E and some class A and D markers; (3) a second region including class A, class D, and non-classed markers; and (4) a terminal region that recombined completely with sex. This latter region was found in the C2 map only.

The markers for regions 1 and 3 were mapped on the X or Y parental chromosomes (Table 5). The recombination rates between the markers mapped in these regions and sex ranged from 0.11 to 0.3. In C1, regions 1 and 3 included one class A marker (5\_254) and one class D marker (8\_164), respectively. The latter marked a fragment common to an X chromosome of the female parent and the Y chromosome (Table 2). In C2, seven markers were mapped in regions 1 and 3. Two of these markers belonged to class A (6\_63 and 1\_318). The class A markers mapped in regions 1 and 3 in C1 or C2 were absent

from some female progenies, indicating recombination with sex in the male parent. In addition to class A markers, regions 1 and 3 included six markers with segregation ratios similar to those of class A markers, but significantly different from the segregation pattern expected for class A (2\_299 in C1; 2\_79, 1\_204, 1\_194, 1\_274 and 1\_364 in C2). The segregation pattern of these markers could be due to distorted segregation, resulting from the preferable inheritance of one of the homologous chromosomes. Among these markers, 2\_299 was detected as a class A marker in C2, supporting the occurrence of distorted segregation in C1 for this marker.

Region 2 included the ‘sex’ phenotypic marker in addition to markers derived from the parental X or Y chromosomes. The markers mapped in this region had recombination rates with sex ranging from 0 to 0.07. In C2, all class E markers except one completely segregated with sex (Table 5). In contrast, three to five recombinants between class E and sex were found in C1, and one class E marker (5\_85) was found in a female plant in C1. Considering both dioecious maps together, all the markers assigned to class E in the present study showed at least one recombinant with sex. In C2, the class A marker 6\_255 and the class D marker 2\_289 were mapped at the sex locus. The





**Fig. 1** Integration of three linkage maps (C1, C2, and UF) showing markers putatively located on sex chromosomes in hemp and location of QTLs for sex expression. C1 and C2 maps were derived from a dioecious population (‘Carmagnola’ × ‘Carmagnola’) and included a ‘sex’ phenotypic marker in C1-1 and C2-1, respectively. UF map was derived from a monoecious population (‘Uso 31’ × ‘Fedora 17’). Name of segregating population and number of co-segregation group are shown above co-segregation groups. Markers are labelled according to number of primer combination (Table 1) and molecular weight of corresponding DNA fragment. Cross type is indicated by different fonts: normal font for markers heterozygous in female parent only, italic for markers heterozygous in male parent only, and bold for markers heterozygous in

both parents. Font colour indicates class of sex-linked markers (Table 2). Kosambi map function was used. LOD and maximum recombination fraction were 3.5 and 0.5, respectively. Coloured lines indicate regions including a putative QTL as revealed by interval mapping (model 1) with  $-\log_{10}(p\text{-value}) \geq 2$  for six variables characterizing sex expression (Table 3). Arrows indicate most likely location of QTL as identified by multi-QTL analysis (model 2). Asterisk indicates modification of scale of QTL effects due to very high likelihood ratios in C1-1 and C2-1: arrows in legend refer to  $-\log_{10}(p\text{-value})$  of 20, 15, and 10 instead of 4, 3, and 2, respectively. Double asterisk indicates markers retained from the cross-population analysis (Table 8)

marker 6\_255 accounted for a fragment mapped on X parental chromosomes, while the marker 2\_289 indicated the presence of a fragment common to an X of the female parent and the Y chromosome of the male parent.

Region 4 was determined by two markers mapped on the X parental chromosomes at the extremity of the sex CG in the C2 map (1\_284 and 6\_336; Table 5). Both markers completely recombined with sex, with recombination rates of 0.41 and 0.50 with sex,

**Table 5** Structure of co-segregation groups (CGs) including the ‘sex’ phenotypic marker in the dioecious maps C1-1 and C2-1. Region in CG<sup>1</sup>, marker name, position along CG, recombination rate with sex, marker class (Table 2), number of male and female progenies according to presence (M)/absence (m) of marker, and parental genotypes

Map—CG	Region <sup>a</sup>	Marker <sup>b</sup>	Position (cM)	Rec. rate with sex <sup>c</sup>	Class	Female progenies		Male progenies		Parents	
						M	m	M	M	X  X	X  Y
C1-1	1	<b>2_299</b>	0	0.17	–	36	7	9	25	a  o	a  o
		<i>5_254</i>	25	0.11	A	39	4	10	24	a  o	a  o
	2	<b>Sex</b>	35.6		–	0	0	34	0	o  o	o  a
		<b>4_276</b>	41	0.05	E	0	43	30	4	o  o	o  a
		<b>5_323</b>	42.3	0.04	E	0	43	31	3	o  o	o  a
		<b>7_118</b>	43.6	0.05	E	0	43	30	4	o  o	o  a
		<b>4_110</b>	45	0.06	–	0	43	29	5	o  o	o  a
		<b>5_85</b>	49	0.05	E	1	42	31	3	o  o	o  a
	3	<i>8_164</i>	54.5	0.19	D	23	20	30	4	a  o	o  a
	C2-1	1	<i>2_79</i>	0	0.25	–	37	11	7	21	a  o
<i>6_63</i>			25.4	0.16	A	42	6	9	19	a  o	a  o
2		<b>2_299</b>	48.3	0.04	A	47	1	15	13	a  o	a  o
		<i>2_289</i>	70.1	0	D	29	19	28	0	a  o	o  a
		<b>6_255</b>	70.1	0	A	48	0	18	10	o  a	a  o
		<b>Sex</b>	70.1		–	0	48	28	0	o  o	o  a
		<b>4_276</b>	70.1	0	E	0	48	28	0	o  o	o  a
		<b>5_323</b>	70.1	0	E	0	48	28	0	o  o	o  a
		<b>7_118</b>	70.1	0	E	0	48	28	0	o  o	o  a
		<b>4_110</b>	70.1	0	E	0	48	28	0	o  o	o  a
		<b>5_85</b>	70.1	0	E	0	48	28	0	o  o	o  a
		<i>6_84</i>	71.7	0.01	E	0	48	27	1	o  o	o  a
<i>1_220</i>		89.5	0.07	A	47	1	20	8	o  a	a  o	
3		<i>1_204</i>	99	0.23	–	45	3	23	5	o  a	a  o
		<i>1_194</i>	109.3	0.22	–	46	2	25	3	o  a	a  o
		<i>1_274</i>	117.3	0.25	–	45	3	24	4	o  a	a  o
		<i>1_364</i>	122.9	0.3	–	44	4	24	4	o  a	a  o
		<i>1_318</i>	143.6	0.18	A	45	3	20	8	o  a	a  o
4		<i>1_284</i>	151.2	0.41	–	32	16	15	13	o  o	a  o
		<i>6_336</i>	166.3	0.5	–	34	14	19	9	a  o	a  o

<sup>a</sup> Regions were determined based on recombination rate of markers with sex

<sup>b</sup> Markers in bold were mapped in both C1 and C2; markers in italics were mapped in UF and in C1 or C2

<sup>c</sup> Recombination rate with sex as estimated by *rf:2pts* function of OneMap (Margarido et al. 2007)

respectively, thus indicating the presence of a pseudoautosomal region. The marker 1\_284 was mapped in the sex CG by linkage in the coupling phase with the class A marker 1\_318 in the male parent. Thus, it was mapped onto the X chromosome of the male parent. The second marker (6\_336) was linked in the coupling phase with marker 1\_284 in the male parent and thus, also mapped onto its X chromosome.

Identification of CGs putatively located on sex chromosomes

Different CGs putatively located on the sex chromosomes were identified in each map through the presence of markers mapped in the sex CG of the dioecious maps. These CGs were referred to as putative sex CGs. Five markers mapped in the sex

**Table 6** Structure of co-segregation groups (CGs) putatively located on sex chromosomes in the dioecious maps C1 and C2. Marker name, position along CG, marker class (Table 2), number of male and female progenies according to presence (M)/absence (m) of marker, and parental genotypes

Map—CG	Marker <sup>a</sup>	Position (cM)	Class <sup>b</sup>	Female progenies		Male progenies		Parents	
				M	m	M	m	Female	Male
C1-2	8_350	0	B*	17	26	13	21	o l l a	o l l o
	6_90	41	B*	20	23	16	18	o l l a	o l l o
	6_243	66.4	B*	21	22	15	19	o l l a	o l l o
	6_189	75.6	B*	22	21	19	15	o l l a	o l l o
	6_202	84.8	B*	22	21	20	14	o l l a	o l l o
	6_253	96.7	B*	27	16	22	12	o l l a	o l l o
	6_249	114.2	B*	19	24	17	17	o l l a	o l l o
	<b>6_255</b>	126.2	B*	15	28	14	20	o l l a	o l l o
	<b>6_337</b>	145.2	–	12	31	11	23	o l l a	o l l o
	6_49	179.7	B*	16	27	14	20	o l l a	o l l o
	<b>6_72</b>	223.2	B*	27	16	21	13	a l l o	o l l o
	<b>6_111</b>	231	–	38	5	32	2	a l l o	o l l a
	5_74	253	–	32	11	20	14	a l l o	o l l o
C1-3	<i>5_71</i>	0	B*	28	15	21	13	a l l o	o l l o
	<b>5_160</b>	6.1	–	42	1	25	9	a l l o	a l l o
C2-2	7_142	0	B*	31	17	19	9	o l l a	o l l o
	7_217	29.3	–	25	23	21	7	o l l a	o l l o
	5_161	33.1	–	34	14	26	2	o l l a	o l l a
C2-3	<b>6_72</b>	49	B*	28	20	10	18	a l l o	o l l o
	<b>5_160</b>	0	–	37	11	18	10	a l l o	a l l o
	2_136	5	–	35	13	18	10	a l l o	a l l o
C2-4	7_130	5	D	31	17	24	4	a l l o	o l l a
	5_175	0	B*	21	27	10	18	a l l o	o l l o
	4_170	9	–	35	13	21	7	a l l o	o l l a
C2-5	2_66	22.8	–	37	11	28	0	a l l o	o l l a
	4_175	0	B*	23	25	18	10	a l l o	o l l o
	4_140	9.3	B*	21	27	11	17	o l l a	o l l o
C2-5	2_158	11.4	–	32	16	18	10	o l l a	o l l a
	6_230	13.5	–	31	17	19	9	o l l a	o l l a
	2_271	22	B*	16	32	11	17	o l l a	o l l o
	3_111	49.5	B*	19	29	13	15	a l l o	o l l o
	6_69	90.8	–	33	15	18	10	o l l o	a l l o
C2-6	1_42	0	–	39	9	24	4	a l l o	a l l o
	1_62	19.2	–	45	3	26	2	a l l o	a l l o
	1_46	26.4	–	46	2	25	3	a l l o	a l l o
	1_72	38	–	44	4	25	3	a l l o	a l l o
	<i>1_60</i>	47.7	–	46	2	26	2	a l l o	a l l o
	1_338	66	–	40	8	25	3	a l l o	a l l o

<sup>a</sup> Markers in bold were mapped in both C1 and C2; markers in italics were mapped in C2 and UF

<sup>b</sup> B\* markers are putative class B markers: they responded to cross type and segregation criteria of class B but were not linked to class A or D markers

**Table 7** QTL analysis for sex-expression related variables in dioecious (C1 and C2) and monoecious (UF) hemp populations

Pop.	Variable <sup>a</sup>	CG <sup>b</sup>	Marker <sup>c</sup>		QTL					
			Name	Position (cM)	Position (cM)	Effect	Estimate ± SE <sup>d</sup>	r <sup>2e</sup>		
C1	Sex	1	2_299	0	2	αP	0.15 ± 0.06**	23.6***		
						αQ	−0.34 ± 0.06***	50.5***		
			δPQ	−0.17 ± 0.06**	–					
	%F-♂	1	5_323	42.3	42.3	αQ	−0.38 ± 0.03***	85.2***		
						αP	8.61 ± 2.72**			
			2	6_337	145.2	143	αP	−4.26 ± 1.36**		
	%M-♀	1	2_299	0	5	αP	11.31 ± 2.3***			
						αQ	−19.33 ± 3.07***			
			2	6_243	66.4	66	αP	−2.35 ± 1*		
				6_253	96.7	97	αP	3.01 ± 1.01**		
C2	Sex	1	2_299	48.3	48.3	αQ	−0.41 ± 0.05***	75.6***		
						δPQ	−0.13 ± 0.04**	–		
			1_220	89.5	89.5	αQ	−0.29 ± 0.05***	73.2***		
			δPQ	0.12 ± 0.05*	–					
			1_318	143.6	144	αQ	0.05 ± 0.02*	6.2*		
	δPQ	−0.09 ± 0.04*	–							
	%F-♂	6	1_62	19.2	19.2	αP	23.29 ± 6.29***			
						αP	−24.34 ± 9.52*			
			1_72	38	32	αP	−24.34 ± 9.52*			
			UF	%MDinter	3	1_106	113.7	113	αQ	−2.72 ± 0.9**
1					4_121	0	22	δPQ	−3.13 ± 0.94***	
log_k	1	6_215	109.7	109.7	δPQ	0.35 ± 0.15*				
		3	1_149	164.0	151	αP	−0.85 ± 0.4*			
		δPQ	1.89 ± 0.5***							
log_NDm	1	4_241	102.7	103	αQ	0.59 ± 0.18**				

<sup>a</sup> For details, see Table 3

<sup>b</sup> In sex CGs, C1-1 and C2-1 (Table 5), αP effects were due to polymorphism between X chromosomes in the female parent, and αQ effects to polymorphism between X and Y chromosomes in the male parent

<sup>c</sup> Name and position of closest marker to QTL

<sup>d</sup> Effect, standard error, and significance of QTLs estimated from model 2

<sup>e</sup> Proportion of phenotypic variation explained by additive QTL effects computed by model 1 for sex

\*  $P < 0.05$ ; \*\*  $P < 0.01$ ; \*\*\*  $P < 0.001$

CG of the C2 map were also found in the UF population (Table 5). Two of them (6\_255 and 2\_289) were mapped at the sex locus, and one of them (6\_336) was mapped in the pseudoautosomal region of the sex CG. These five markers were mapped in three distinct putative sex CGs in UF (UF-1, 2 and 3; Fig. 1). The mapping of 6\_255 and 6\_336 in the same CG (UF-1) and the co-segregation of 2\_289 and 1\_194 (UF-2) in the UF map supported the presence of 6\_336 and 1\_194 on the sex chromosomes. These two markers

showed relatively high recombination rates with sex in C2 (Table 5).

Two putative sex CGs were identified in the C1 map (Table 6). C1-2 included a marker that was linked to the sex and assigned to class A in C2 (6\_255). Among the 13 markers linked in C1-2, 12 consisted of an allele derived from the female parent only and 10 were putative class B markers, being heterozygous in the female parent only and segregating 1:1 in both male and female progenies. One of them (6\_72) was also a

putative class B marker in C2. The second CG putatively located on the sex chromosomes in C1 (C1-3) was identified through the presence of one marker common to a putative sex CG in UF (5\_71, present in both C1-3 and UF-1). This marker was a putative class B marker in C1 and was linked to a class D marker (7\_130) in C2-3.

Five putative sex CGs were identified in the C2 map (Table 6). In C2-4, one marker heterozygous in both parents (2\_66) was completely linked to the sex in the male parent, but its segregation in the female progenies did not correspond to the expected 1:1 ratio of class D markers (Table 2). C2-2, -4, and -5 included two, one, and four putative class B markers, respectively. C2-3 included a marker for which the cross type and segregation ratios in the female and male progenies corresponded to a class D marker at  $\alpha = 0.01$  (7\_130).

QTL analysis

In total, 336, 440, and 515 points were evaluated in the C1, C2, and UF maps, respectively. The search for QTLs by interval mapping and multi-QTL analysis led to the identification of five distinct QTLs for sex expression in each map. The total number of QTLs per population was the sum of the individual QTLs detected for different phenotypic variables (Table 7; Fig. 1).

(i) Dioecious populations (C1 and C2)

Because of colinearity, the multi-QTL analysis for sex was performed after discarding the sex locus from the multi-QTL model (model 2). In C1, two QTLs were found for sex, both mapped in the sex CG (C1-1). One of them was mapped close to a class E, male-associated, marker (5\_323; Tables 5, 7) and explained

**Table 8** Cross-population analysis of sex expression: markers that segregated with sex in the dioecious populations (C1 and C2) and accounted for a putative QTL in the monoecious

population (UF), or accounted for a putative QTL in a dioecious population (C1 or C2) and were associated with variation of sex expression in UF

CG no.			Marker	Dioecious populations					Monoecious population		Variables affected by the marker locus <sup>e</sup>
C1	C2	UF		Frequency in ♀-♂ <sup>a</sup>		Test of the effects <sup>b</sup>			Variable <sup>c</sup>	Estimate ± SE <sup>d</sup>	
				C1	C2	M	Pop	M × Pop			
–	–	1	4_121	84–94	88–100	*	NS	NS	–	–	Log_k (UF)
–	4	1	2_66	77–100**	77–100**	***	NS	NS	log_NDm	–0.52 ± 0.18**	Sex (C2), Log_NDm (UF)
2	1	1	6_255	35–41	100–64***	***	NS	***	mMD	0.16 ± 0.08*	Sex (C2),
									%MD1	–4.27 ± 1.96*	Log_NDm
									log_k	0.28 ± 0.13*	(UF)
									log_NDm	–0.53 ± 0.18**	
–	5	1	2_271	23–32	33–39	NS	NS	NS	mMD	–0.21 ± 0.09	%F-♂ (C2)
									%MD1	5.06 ± 2.05*	
									log_NDm	0.41 ± 0.2 *	
									log_Mmax	–0.16 ± 0.07*	

<sup>a</sup> Frequency of marker in female progenies and in male progenies (%)

<sup>b</sup> Consistency of segregation pattern of marker with sex across both dioecious populations: effect of marker (M), population (Pop) and their interaction on sex obtained using model 3

<sup>c</sup> For details, see Table 3

<sup>d</sup> Estimate, standard error, and significance of marker effect on sex-expression variables in monoecious population obtained using model 3

<sup>e</sup> Phenotypic variables showing variation associated with genetic predictor at marker locus, as revealed by model 1

NS  $P > 0.05$ ; \*  $P < 0.05$ ; \*\*  $P < 0.01$ ; \*\*\*  $P < 0.001$

a large proportion of the sex variation (85.2 %). The second QTL for sex was mapped at the locus of a class A marker (2\_299), which is found on X chromosomes in both the female and male parent, and had a dominance effect. In C2, three QTLs with both additive and dominance effects were detected for sex. All of them were mapped in the sex CG (C2-1) close to markers mapped on X parental chromosomes. One of them was common to C1 (2\_299). Large proportions of phenotypic variation (> 73 %) were explained by the two closest QTLs to the sex locus (2\_299 and 1\_220).

In C1, two QTLs were identified for the percentage of nodes bearing female flowers in male plants (% $F$ -♂), and three QTLs for the percentage of nodes bearing male flowers in female plants (% $M$ -♀). One of them (2\_299) was a QTL for both % $F$ -♂ and % $M$ -♀ in addition to sex. This QTL had a positive additive effect on % $F$ -♂ and a negative effect on % $M$ -♀, both due to the male parent ( $\alpha_0$ ). Therefore, this QTL could include feminising genetic factors located on the male-parent X chromosome (Table 7). In C2, two QTLs were identified for % $F$ -♂ in a putative sex CG (C2-6), and none for % $M$ -♀. Except for 2\_299 in C1, which had both additive and dominance effects, all of the QTLs detected for % $F$ -♂ and % $M$ -♀ in C1 and C2 had additive effects only. These QTLs were located in putative sex CGs.

#### (ii) Monoecious population (UF)

Five QTLs were found for variables related to sex expression (Table 7) in UF. All except one were associated with variation in structure variables—i.e., the parameters of the logistic function describing the monoecy degree along the stem (Faux and Bertin 2014). Two QTLs had additive effects only on sex expression: 1\_106 (UF-3) on the percentage of nodes with intermediate monoecy degree (% $MD_{inter}$ ), and 4\_241 (UF-1) on the parameter accounting for the node at which there is maximum variation in monoecy degree along the stem ( $log\_ND_m$ ). One of them (4\_241) was closely linked to marker 6\_255, which was itself mapped at the sex locus in C2 (Fig. 1). The three other QTLs had both additive and dominance effects (1\_149 in UF-3), or only dominance effects on sex expression (4\_121 and 6\_215 in UF-1). One of these QTLs (6\_215) was also identified close to marker 6\_255.

#### Cross-population analysis

The cross-population analysis identified three markers that accounted for a putative QTL in UF and segregated with sex in both C1 and C2 (4\_121, 2\_66 and 6\_255, showing a significant ‘M’ effect in Table 8), and one marker that accounted for a putative QTL in C2 and was associated with variation of sex expression in UF (2\_271). All of these four markers were detected in the three populations; however, all of them were mapped in UF only (UF-1; Fig. 1). The identification in UF-1 of markers that segregated with sex in both C1 and C2 supported the presence of the CG on the sex chromosomes of the monoecious population. Three of these markers were mapped at positions that were, by homology, close to the sex locus (2\_66, 6\_255, and 2\_271; Fig. 1).

The first marker (4\_121) was more frequent in male plants across both dioecious populations and accounted for a QTL for the parameter determining the curvature of the logistic function ( $log\_k$ ) in UF. The second marker (2\_66) was present in all male plants and in 77 % of female plants in both C1 and C2, and it was associated with variation in  $log\_ND_m$  in UF. The third marker (6\_255) was mapped on the X chromosomes at the sex locus in C2. It was a class A marker in C2 and a putative class B marker in C1 (Tables 5, 6). In UF, this marker was associated with feminised phenotypes (positive effect on the mean monoecy degree,  $mMD$ , and negative effect on the percentage of highly masculinised nodes along the stem, % $MDI$ ), an increased  $log\_k$ , and a decreased  $log\_ND_m$ . The fourth marker (2\_271) was a putative class B marker in C2 (Table 6). This marker was associated with variation in four sex-expression variables in UF (Table 8).

#### Discussion

In this study, we focused on the role of the sex chromosomes in the genetic determinism of sex expression in both dioecious and monoecious hemp. First, we constructed three genetic maps of the sex chromosomes in hemp: two derived from a cross between female and male plants of the dioecious cultivar ‘Carmagnola’, and one derived from a cross between the monoecious cultivars ‘Fedora 17’ and ‘Uso 31’. The genetic determinism of sex expression

was then investigated using a QTL analysis. Because of the high variability of the sexual phenotype in hemp, we searched for QTLs based on distinct variables that dissect the variability of sex expression (Table 3; Faux and Bertin 2014). To our knowledge, this is the first study to integrate linkage maps from both dioecious and monoecious hemp, and the first report of a QTL analysis in hemp.

#### AFLP markers

The percentage of markers scored simultaneously in both dioecious populations was relatively high (81 %), as expected from their common ‘Carmagnola’ origin, although the percentage of markers scored simultaneously in one dioecious population and in the monoecious population was also large (67 % for C1 and UF, and 68 % for C2 and UF). This was consistent with the structure of the genetic diversity reported in hemp by Forapani et al. (2001), who concluded to the existence of a widely shared gene pool with limited genetic separations among groups.

The percentage of segregating markers was the lowest in C2 (44 %), followed by UF (52 %) and C1 (67 %). These values were relatively high compared to the percentage of RAPD-segregating markers found in a hemp progeny obtained from a cross between a ‘Carmagnola’ female plant and a monoecious plant (39.1 %) by Carboni et al. (2000). The difference in percentage of segregating markers found between UF and C1 can be explained by the narrower genetic basis of monoecious hemp, as a result of its ability to self-pollinate (Bocsa and Karus 1998) and the selection pressure needed to maintain the monoecious trait (Forapani et al. 2001; Mandolino and Carboni 2004). However, the percentage of segregating markers found in the dioecious population C2 was particularly low. It indicated the presence of a relatively high proportion of markers at the dominant homozygous stage (‘aa’) in one or both parents, resulting in a relatively high proportion of not segregating loci. This could be attributed to the relatively large genetic diversity of hemp, which has been characterized by high proportions of polymorphic markers, including within the cultivar ‘Carmagnola’ (Forapani et al. 2001).

The number of markers that segregated independently in all three populations (19 among 319 independently segregating markers) and the number of

markers shared by the three maps (5 among 225 mapped markers) were low in contrast with the relatively large number of markers simultaneously scored in all three populations (287 among 480 scored markers). Similarly, Waugh et al. (1997) identified only eight AFLP markers that segregated in three populations of barley among totals of 234, 194 and 376 mapped AFLP markers, while Hoarau et al. (2001) used a relatively low number of anchoring markers (45 AFLPs) to join two maps that included 887 AFLP and 408 RFLP markers in sugarcane.

In particular, relatively few markers segregated independently in both ‘Carmagnola’ populations (75 markers) and were finally shared by both ‘Carmagnola’ maps (22 markers). This can be explained by two effects. First, 47 % of the markers that segregated independently in C1 were fixed in C2. This indicated the presence of markers at the dominant homozygous stage in a given parent of the C2 population but not in the parents of the C1 population, resulting in markers that are fixed in C2 but segregating in C1, and consequently in a low number of markers segregating in both populations. Second, the percentage of markers heterozygous in both parents among the independently segregating markers was relatively high in both C1 and C2 (63 and 64 %; Table 4). Such bi-allelic markers heterozygous in both parents and segregating 3:1 in the offspring are hereafter referred to as C.8 markers according to the notation of Wu et al. (2002). According to Maliepaard et al. (1997), the presence of a high proportion of C.8 markers decreases the probability of detecting linkages between them, especially when the population size is small. The relatively high percentages of C.8 markers in C1 and C2 and the relatively small size of these populations (77 and 76 progenies) might have prevented the detection of linkages among some markers, thereby resulting in a relatively low number of markers that are linked in both C1 and C2. Nevertheless, the sizes of the ‘Carmagnola’ populations used in this study were equivalent to those of populations used in mapping derived from crosses between male and female plants of dioecious hemp accessions, i.e., 66 (Mandolino and Ranalli 2002) and 80 individuals (Peil et al. 2003).

#### Linkage maps

The three maps of the sex chromosomes of hemp were unsaturated, as reflected by the mapping of several co-

segregation groups that included few markers (Fig. 1). As a result, except for the sex CGs (C1-1 and C2-1), the location of the CGs on the sex chromosomes could not be ascertained. Nevertheless, the mapping of putative class B markers or a class D marker in each of the putative sex CGs from the C1 and C2 map except for C2-6 supported the presence of these CGs on the sex chromosomes.

The integration of the three maps revealed different inconsistencies with respect to marker linkages. For example, there were linkages between markers mapped together in the sex CG in C2 but in different putative sex CGs in UF. Also, compared with the C1 and UF maps, the C2 map had more CGs (11, 16, and 10 CGs, among which 3, 6, and 3 were putatively located on the sex chromosomes in C1, C2, and UF, respectively). These inconsistencies were attributed to the relatively low accuracy of the recombination fraction estimates and the low number of markers shared by the present maps, including between the two ‘Carmagnola’ maps. According to Maliepaard et al. (1997), the accuracy of the recombination fraction estimates in the C2 map was negatively affected by the relatively high proportion of C.8 markers (67 % against 56 % in C1 and 41 % in UF), and, in both ‘Carmagnola’ maps, by their relatively small population sizes. Indeed, the high uncertainty for the linkages among C.8 markers can prevent the detection of linkages between some markers, which would in turn result in a higher number of small CGs in C2, as well as make it difficult to establish the correct order of markers. As discussed above, the numbers of markers shared among the present maps were relatively low, and this makes the integration of the maps more difficult.

The consecutive mapping of markers derived from the same primer combination in some CGs (C1-2, C2-1, and C2-6) could be due to the presence of repetitive DNA sequences. In hemp, Peil et al. (2003) observed clusters of AFLP markers along linkage groups. Similarly, Rogers et al. (2007) reported a non-random distribution of AFLP loci amplified using different selective primer combinations across linkage groups in whitefish (*Coregonus clupeaformis*). They raised the possibility that repetitive DNA in the genome, which is partially due to transposable elements, could result in the occurrence of repetitive AFLP sequences. In a study on *Asparagus officinalis*, Reamon-Büttner et al. (1999) observed clusters of hybridisation signals

on chromosomes, suggesting that the AFLP fragments were parts of repetitive sequence families. According to Sakamoto et al. (2005), multiple sequences encoding retrotransposable elements are ubiquitous in the hemp genome.

#### Identification of sex-linked markers and CGs putatively located on sex chromosomes of the monoecious population

Four male-specific markers detected in the present study had sizes similar to male-specific AFLP markers reported in hemp: the markers 4\_251, 4\_276, and 5\_323 correspond to ACC\*AAG250, ACC\*AAG275, and ACC\*AGA323, respectively, in the study of Flachowsky et al. (2001).

The detection of sex-linked markers allowed (i) the characterisation of the structure of the sex CG in the dioecious maps and (ii) the identification of markers and CGs putatively located on the sex chromosomes in the monoecious map. Like Peil et al. (2003), we identified a region including markers of classes A, D, and E that were closely linked to sex (region 2), as well as a region including markers that completely recombined with sex at an extremity of the sex CG (region 4 in the C2 map). These observations confirmed the presence of common fragments between X and Y chromosomes—represented by class D—and the existence of a pseudoautosomal region in the sex chromosomes of hemp, as reported by Peil et al. (2003). However, our results differed from those of Peil et al. (2003) in that we observed higher recombination rates between the markers mapped in each of the four regions and sex. First, one to five recombinants were observed between each of our class E markers and sex in C1 or C2 (Table 5). In contrast, Peil et al. (2003) detected only one recombinant with sex and therefore assumed that their class E markers were located in a non-pairing portion of the Y chromosome. The recombinations observed between our class E markers and sex suggested that they were located in a region where pairing between X and Y chromosomes can occur. Second, the recombination rates found here between the pseudoautosomal markers and sex ( $\geq 0.41$ ) were relatively high compared with the values of 0.25 and 0.27 in hemp reported by Peil et al. (2003) and Rode et al. (2005), respectively. Third, two additional regions were distinguished in the sex CGs (regions 1 and 3; Table 5). Without



considering the distorted segregations of the markers, the recombination rates with sex observed in regions 1 and 3 ranged from 0.11 to 0.19. The recombination rates observed here in the sex CG (Table 5) suggest that the X and Y chromosomes of hemp recombine with each other between the sex locus and the pseudoautosomal region. The existence of recombination between the X and Y chromosomes in the male parent is supported by the pairing of both chromosomes at the short arm of the Y chromosome reported by Sakamoto et al. (2000).

As well as identifying sex-linked markers in dioecious hemp, we identified sex-linked markers in monoecious hemp. Five markers mapped in the sex CG of C2 were detected in UF (Table 5). Four of these were mapped on X chromosomes, including one in the pseudoautosomal region, and one was mapped on both X and Y chromosomes (class D). Since monoecious hemp has XX chromosomes (Menzel 1964; Faux et al. 2014), the mapping of pseudoautosomal and class D markers in monoecious hemp suggests the presence of homologous fragments between the Y chromosome of dioecious hemp and the X chromosomes of monoecious hemp.

#### Determinism of sex expression in dioecious hemp

Five QTLs associated with sex expression were identified in both C1 and C2 (Table 7). One of them was detected in both maps (2\_299). Large proportions of phenotypic variation were explained by the QTLs for sex (from 73.2 to 85.2 % for the closest QTL to sex), which was linked to the presence of heteromorphic sex chromosomes. In addition to QTLs for sex, QTLs were identified for the percentage of nodes bearing flowers of the sex opposite to the main one (%*F*-♂ in male plants and %*M*-♀ in female plants). The identification of QTLs for those variables supported previous assumptions that there is a genetic basis for the production of flowers of the sex opposite to the main one in dioecious hemp [Borthwick and Scully 1954; Grassi and de Meijer (pers. comm.) cited by Moliterni et al. 2004].

According to our results, the role of sex chromosomes in the determinism of sex expression in dioecious hemp could be as follows. The region surrounding the sex locus (region 2 in Table 5) would include genetic factors involved in the differentiation of the male plants, as suggested by the identification of

a QTL with an additive effect due to the male parent in C1 (5\_323; Table 7; Fig. 1). In contrast to region 2, regions 1 and 3 of the sex CGs would not be involved in the differentiation of male plants given the higher recombination rates with the sex locus and the presence of only one QTL for sex with a dominance effect and a low  $r^2$  (1\_318 in C2; Tables 5, 7). Instead, these regions would include genetic factors involved in the production of flowers of the sex opposite to the main one, as suggested by the QTL for both %*F*-♂ and %*M*-♀ identified in C1 (2\_299; Table 7). No QTL was detected in the terminal pseudoautosomal region of the sex CG in C2. However, beyond the four regions defined in the sex CGs, our results suggest that additional regions of the sex chromosomes carry genetic factors involved in the determinism of sex expression. Indeed, QTLs for %*F*-♂ and %*M*-♀ were mapped in putative sex CGs (6\_337, 6\_243 and 6\_253 in C1-2, and 1\_62 and 1\_72 in C2-6; Table 7; Fig. 1). These QTLs had additive effects due to the female parent and would therefore be located on X chromosomes only.

#### Determinism of sex expression in monoecious hemp

In this study, we identified five distinct QTLs associated with sex expression in UF (Table 7), and four markers that were associated with variation of sex expression in UF and segregated with sex, or accounted for a putative QTL for sex expression, in C1 or C2 (Table 8). Two of these QTLs (4\_241 and 6\_215) and three of these markers (2\_66, 6\_255, and 2\_271) were mapped in a region homologous to the sex-locus region of the dioecious maps. These results suggested that genetic factors involved in the determinism of sex expression in monoecious hemp exist on the sex chromosomes, and thus on X chromosomes. In addition, these results suggested that genetic factors with quantitative effects on sex expression would be closely linked to the sex locus on the X chromosomes in monoecious hemp.

The number of QTLs detected for each phenotypic variable related to sex expression in UF ranged from zero to three (Tables 3, 7). In total, the synthesis variables allowed the detection of one QTL (1\_160 for %*MDinter*), while the structure variables, which consisted of parameters of a logistic curve describing sex expression as a function of node position

(Table 3), allowed the detection of four distinct QTLs. The higher number of QTLs found using the modelling approach highlights its utility for characterizing the variability of sex expression among monoecious hemp plants.

## General discussion

The identification of QTLs for the quantitative variation of sex expression on the sex chromosomes of hemp and the recombination rates observed between the sex chromosomes contrasted with the situation in *Silene latifolia*, a well-studied dioecious species with heteromorphic sex chromosomes (Filatov et al. 2001). In *S. latifolia*, recombination is absent from most of the Y chromosome (Charlesworth 2002), while the genetic basis of sex determination is strong, and there is little evidence for lability or environmental effects (Ainsworth 2000). According to Charlesworth et al. (2005), the suppression of recombination between the sex chromosomes in dioecious species results from the presence of sex-determining genes and the evolution of Y-linked genes that benefit male but not female functions, both effects resulting in selection against recombinants. In hemp, it is possible that the recombination rates observed between the sex chromosomes allow the exchange of genetic factors that affect the production of flowers of a given sex. Compared with *Silene*, in hemp, the individuals with recombined sex chromosomes would have a relatively high adaptive value, as supported by the diversity of intersexual forms existing in the species.

According to Dellaporta and Calderon-Urrea (1993), the QTLs identified for sex expression in the present study could include genetic factors that regulate programs of sexuality through a signal transduction mechanism that modifies endogenous hormonal levels. Indeed, the formation of male and female generative organs in hemp may be associated with an increased demand for gibberellin and auxins, respectively (Galoch 1980). More recently, cDNA-AFLP fragments differentially expressed in female and male apices were identified. These fragments showed similarities to a Rac-GTP binding protein that plays a signalling role in auxin-regulated gene expression in *Arabidopsis* (Moliterni et al. 2004). These studies suggested that sex expression in hemp could be related to the presence of hormonal gradients

along the stem. This could provide a physiological explanation for the model parameters used in this study to characterize the sex expression in monoecious hemp. Further studies combining physiological and QTL approaches are needed to test this assumption.

## Conclusions

In this study, we identified QTLs associated with the quantitative variation in sex expression on the sex chromosomes of both dioecious and monoecious hemp, despite the high environmental sensitivity of this trait. For this purpose, we constructed three AFLP linkage maps of the sex chromosomes in hemp. Two maps were derived from dioecious populations and one was derived from a monoecious population. Although these maps were unsaturated, they allowed us to identify sex-linked markers and, by homology, co-segregation groups putatively located on the sex chromosomes of dioecious and monoecious hemp.

The main advances can be summarized as follows. First, the X and Y chromosomes of dioecious hemp would recombine with each other between the sex locus and the pseudoautosomal region. Second, the X chromosomes of monoecious hemp would include fragments homologous with both the X and Y chromosomes of dioecious hemp. Third, the sex chromosomes of both dioecious and monoecious hemp would include genetic factors associated with quantitative variations in sex expression. Fourth, the higher number of QTLs detected for the parameters of a logistic curve modelling sex expression as a function of the node position along the stem supported the relevance of this approach for characterizing the variability of sex expression among monoecious hemp plants. Finally, the results of this study suggested that it would be relevant to conduct further research on the genetic determinism of sex expression in hemp using a quantitative approach.

**Acknowledgments** We thank Dr. H. Flachowsky and Dr. A. Peil (Julius Kühn-Institut, Germany) for sharing their AFLP protocol, V. Van Hese and P. Van Thorre (Université Catholique de Louvain) for valuable help with laboratory work and with the cultivation of hemp plants, respectively. We thank the platform SMCS (Support in Methodology and Statistical Computation) of the Université Catholique de Louvain for statistical advice. This work was supported by the Fonds de la Recherche Scientifique [Grant no. 1.5.279.08].

## References

- Ainsworth C (2000) Boys and girls come out to play: the molecular biology of dioecious plants. *Ann Bot* 86:211–221
- Arnoux M (1966a) Influence des facteurs du milieu sur l'expression de la sexualité du chanvre monoïque (*Cannabis sativa* L.). II—Action de la nutrition azotée. *Ann Amélior Plantes* 16:123–134
- Arnoux M (1966b) Influence des facteurs du milieu sur l'expression de la sexualité du chanvre monoïque (*Cannabis sativa* L.). III—Note sur l'interaction entre le cycle photopériodique et la nutrition azotée. *Ann Amélior Plantes* 16:259–262
- Arnoux M, Mathieu G (1969) Influence du milieu sur le phénotype sexuel de descendances F<sub>1</sub> issues du croisement entre types dioïques et monoïque du chanvre (*Cannabis sativa* L.). *Ann Amélior Plantes* 19:53–58
- Beherec O (2000) FNPC hemp breeding and CCPSC seeds production (France). In: Proceedings of “Bioresource Hemp 2000”, 13–16 Sept 2000, Wolfsburg, Germany
- Bocsa I, Karus M (1998) The cultivation of hemp: botany, varieties, cultivation and harvesting. HempTech, Sebastopol
- Boer MP, Wright D, Feng L et al (2007) A mixed-model quantitative trait loci (QTL) analysis for multiple-environment trial data using environmental covariables for QTL-by-environment interactions, with an example in maize. *Genetics* 177:1801–1813
- Borthwick HA, Scully NJ (1954) Photoperiodic responses in hemp. *Bot Gaz* 116:14–29
- Carboni A, Paoletti C, Moliterni VMC, Ranalli P, Mandolino G (2000) Molecular markers as genetic tools for hemp characterization. Proceedings of the “Bioresource Hemp 2000”, 13–16 September 2000, Wolfsburg, Germany
- Charlesworth D (2002) Plant sex determination and sex chromosomes. *Heredity* 88:94–101
- Charlesworth D, Charlesworth B, Marais G (2005) Steps in the evolution of heteromorphic sex chromosomes. *Heredity* 95:118–128
- Dellaporta SL, Calderon-Urrea A (1993) Sex determination in flowering plants. *Plant Cell* 5:1241–1251
- Faux A-M, Bertin P (2014) Modelling approach for the quantitative variation of sex expression in monoecious hemp (*Cannabis sativa* L.). *Plant Breed* 133:679–793
- Faux A-M, Draye X, Lambert R, d'Andrimont R, Raulier P, Bertin P (2013) The relationship of stem and seed yields to flowering phenology and sex expression in monoecious hemp (*Cannabis sativa* L.). *Eur J Agron* 47:11–22
- Faux A-M, Berhin A, Dauguet N, Bertin P (2014) Sex chromosomes and quantitative sex expression in monoecious hemp (*Cannabis sativa* L.). *Euphytica* 196:183–197
- Filatov DA, Laporte V, Vitte C, Charlesworth D (2001) DNA diversity in sex-linked and autosomal genes of the plant species *Silene latifolia* and *Silene dioica*. *Mol Biol Evol* 18:1442–1454
- Flachowsky H, Schumann E, Weber WE, Peil A (2001) Application of AFLP for the detection of sex-specific markers in hemp. *Plant Breed* 120:305–309
- Forapani S, Carboni A, Paoletti C, Moliterni VMC, Ranalli P, Mandolino G (2001) Comparison of hemp varieties using random amplified polymorphic DNA markers. *Crop Sci* 41:1682–1689
- Fournier G, Beherec O (2006) La sélection du chanvre. In: Bouloc P (ed) *Le chanvre industriel, production et utilisations*. France Agricole, Paris, pp 89–110
- Freeman DC, Harper KT, Charnov EL (1980) Sex change in plants: old and new observations and new hypotheses. *Oecologia* 232:222–232
- Galoch E (1980) The hormonal control of sex differentiation in dioecious plants of hemp *Cannabis sativa* 3. The level of phytohormones in male and female plants in the different stages of their development. *Acta Physiol Plant* 2:31–40
- Hoarau JY, Grivet L, Offmann B, D'Hont A, Risterucci AM, Roques D, Glaszmann JC, Grivet L (2001) Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp.) I. Genome mapping with AFLP markers. *Theor Appl Genet* 103:84–97
- Lander ES, Botstein D (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185–199
- Lynch M, Walsh B (1998) *Genetics and analysis of quantitative traits*. Sinauer Associates, Sunderland
- Maliepaard C, Jansen J, van Ooijen JW (1997) Linkage analysis in a full-sib family of an outbreeding plant species: overview and consequences for applications. *Genet Res* 70:237–250
- Malosetti M, Voltas J, Romagosa I et al (2004) Mixed models including environmental covariables for studying QTL. *Euphytica* 37:139–145
- Malosetti M, Visser RGF, Celis-Gamboa C, van Eeuwijk F (2006) QTL methodology for response curves on the basis of non-linear mixed models, with an illustration to senescence in potato. *Theor Appl Genet* 113:288–300
- Malosetti M, Ribaut JM, Vargas M et al (2007) A multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea mays* L.). *Euphytica* 161:241–257
- Mandolino G, Carboni A (2004) Potential of marker-assisted selection in hemp genetic improvement. *Euphytica* 140:107–120
- Mandolino G, Ranalli P (2002) The applications of molecular markers in genetics and breeding of hemp. *J Ind Hemp* 7:7–23
- Margarido GRA, Souza AP, Garcia AAF (2007) OneMap: software for genetic mapping in outcrossing species. *Hereditas* 144:78–79
- Margarido G, Mollinari M, Broman K, Garcia A (2012) Package ‘onemap’: software for constructing genetic maps in experimental crosses: full-sib, RILs, F2 and backcrosses. Version 2.0-3. <http://cran.r-project.org/web/packages/onemap/index.html>
- Menzel MY (1964) Meiotic chromosomes of monoecious Kentucky hemp (*Cannabis sativa*). *Bull Torrey Bot Club* 91:193–205
- Meudt HM, Clarke AC (2007) Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends Plant Sci* 12:106–117

- Moliterni VMC, Cattivelli L, Ranalli P, Mandolino G (2004) The sexual differentiation of *Cannabis sativa* L.: a morphological and molecular study. *Euphytica* 140:95–106
- Mueller U, Wolfenbarger L (1999) AFLP genotyping and fingerprinting. *Trends Ecol Evol* 14:389–394
- Pastina MM, Malosetti M, Gazaffi R et al (2012) A mixed model QTL analysis for sugarcane multiple-harvest-location trial data. *Theor Appl Genet* 124:835–849
- Peil A, Flachowsky H, Schumann E, Weber WE (2003) Sex-linked AFLP markers indicate a pseudoautosomal region in hemp (*Cannabis sativa* L.). *Theor Appl Genet* 107:102–109
- Qi X, Stam P, Lindhout P (1998) Use of locus-specific AFLP markers to construct a high-density molecular map in barley. *Theor Appl Genet* 96:376–384
- Ranalli P (2004) Current status and future scenarios of hemp breeding. *Euphytica* 140:121–131
- Razumova OV, Alexandrov OS, Divashuk MG, Sukhorada TI, Karlov GI (2015) Molecular cytogenetic analysis of monoecious hemp (*Cannabis sativa* L.) cultivars reveals its karyotype variations and sex chromosomes constitution. *Protoplasma*. doi:10.1007/s00709-015-0851-0
- Reamon-Büttner SM, Jung C (2000) AFLP-derived STS markers for the identification of sex in *Asparagus officinalis* L. *Theor Appl Genet* 100:432–438
- Reamon-Büttner SM, Schmidt T, Jung C (1999) AFLPs represent highly repetitive sequences in *Asparagus officinalis* L. *Chromosome Res* 7:297–304
- Rode J, In-Chol K, Saal B, Flachowsky H, Kriese U, Weber WE (2005) Sex-linked SSR markers in hemp. *Plant Breed* 170:167–170
- Rogers SM, Isabel N, Bernatchez L (2007) Linkage maps of the dwarf and normal lake whitefish (*Coregonus clupeaformis*) species complex and their hybrids reveal the genetic architecture of population divergence. *Genetics* 175:375–398
- Roupe van der Voort J, Draaistra J, Gommers FJ (1997) Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. *Mol Gen Genet* 255:438–447
- Sakamoto K, Shimomura K, Komeda Y, Kamada H, Satoh S (1995) A male-associated DNA sequence in a dioecious plant, *Cannabis sativa* L. *Plant Cell Physiol* 36:1549–1554
- Sakamoto K, Ohmido N, Fukui K, Kamada H, Satoh S (2000) Site-specific accumulation of a LINE-like retrotransposon in a sex chromosome of the dioecious plant *Cannabis sativa*. *Plant Mol Biol* 44:723–732
- Sakamoto K, Abe T, Matsuyama T, Yoshida S, Ohmido N, Fukui K, Satoh S (2005) RAPD markers encoding retrotransposable elements are linked to the male sex in *Cannabis sativa* L. *Genome* 48:931–936
- SAS Institute Inc (2012) SAS<sup>®</sup> 9.3 help and documentation. SAS Institute Inc, Cary
- Sengbusch RV (1952) Ein weiterer beitrage zur vererbung des geschlechts bei hanf als grundlage für die züchtung eines monözischen hanfes. *Z Pflanzenzüchtung* 31:319–338
- Struik PC, Amaducci S, Bullard MJ, Stutterheim NC, Venturi G, Cromack HTH (2000) Agronomy of fibre hemp (*Cannabis sativa* L.) in Europe. *Ind Crop Prod* 11:107–118
- Terauchi R, Kahl G (1999) Mapping of the *Dioscorea tokoro* genome: AFLP markers linked to sex. *Genome* 42:752–762
- Truta E, Olteanu N, Surdu S, Zamfirache MM, Oprica L (2007) Some aspects of sex determinism in hemp. *Analele Stiintifice ale Universitatii « Alexandru Ioan Cuza ». Sectiunea Genetica si Biologie Moleculara VIII*:31–39
- van der Werf HMG, Mathijssen EWJM, Haverkort AJ (1996) The potential of hemp (*Cannabis sativa* L.) for sustainable fibre production: a crop physiological appraisal. *Ann Appl Biol* 129:109–123
- Vos P, Hogers R, Bleeker M, Reijmans M, Van De Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* 23:4407–4414
- Waugh R, Bonar N, Baird E, Thomas B, Graner A, Hayes P, Powell W (1997) Homology of AFLP products in three mapping populations of barley. *Mol Gen Genet* 255:311–321
- Westergaard M (1958) The mechanism of sex determination in dioecious flowering plants. *Adv Genet* 9:217–281
- Wu R, Ma CX, Painter I, Zeng ZB (2002) Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing species. *Theor Popul Biol* 61:349–363